

## ТЕХНОЛОГИЯ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ В ПРЕДСКАЗАНИИ ФИЗИКО-ХИМИЧЕСКИХ СВОЙСТВ ОРГАНИЧЕСКИХ МОЛЕКУЛ

Туманов В.Е.<sup>1</sup>, Прохоров А.И.<sup>1</sup>, Соловьева М.Е.<sup>1</sup>

<sup>1</sup>Институт проблем химической физики РАН, Черноголовка, Россия, (142432, Черноголовка, проспект академика Семенова, 1), e-mail: aipro@icp.ac.ru

В статье представлена облачная и многоагентная программно-технологическая архитектура предметно-ориентированной системы научной осведомленности по физической химии радикальных реакций, дано описание его функционала. Электронные коллекции физико-химических свойств органических соединений, объединенные в системе, позволяют ставить и решать задачи извлечения новых знаний о свойствах этих соединений. Приведено описание назначения используемых интеллектуальных агентов в решении задач предсказания физико-химических свойств органических молекул по экспериментальным данным и их место в архитектуре системы. Дан пример разработки интеллектуального программного агента для предсказания реакционной способности молекул нечеткой базой знаний. Впервые на примере реакций фенил радикалов с углеводородами была выполнена попытка идентифицировать зависимость классического потенциального барьера радикальной реакции нечеткой базой знаний, построенной на основе количественных и качественных параметров.

Ключевые слова: облачная архитектура, интеллектуальный программный агент, нечеткая база знаний, искусственная нейронная сеть, кластерный анализ, радикальная реакция, жидкая фаза, реакционная способность, энтальпия реакции, энергия диссоциации связей

## TECHNOLOGY OF INTELLIGENT AGENTS IN THE PREDICTION OF PHYSICAL AND CHEMICAL PROPERTIES OF ORGANIC MOLECULES

Tumanov V.E.<sup>1</sup>, Prokhorov A.I.<sup>1</sup>, Soloveva M.E.<sup>1</sup>

<sup>1</sup>Institute of problems of chemical physics RAS, Chernogolovka, Russia (142432, Chernogolovka, ac. Semenov av., 1), e-mail: aipro@icp.ac.ru

The architecture of cloud computing and multi-agent software technologies for the scientific subject-oriented knowledge system in physical chemistry of radical reactions is presented in article; the description of its functionality is given. Electronic collections of physical and chemical properties of organic compounds which combined in the system allow to set and to achieve objectives of extraction of new knowledge about the properties of these compounds. The destination of intelligent agents to solve the prediction of physical and chemical properties of organic molecules on base of experimental data and their place in the architecture of the system is described. The example of the development of intelligent software agents to predict the reactivity of molecules by means of a fuzzy knowledge base is adduced. First attempt was made to identify the dependence of the classical potential barrier for phenyl radical reactions with hydrocarbons by means of the fuzzy knowledge base which was constructed on the basis of quantitative and qualitative parameters of the radical reactions.

Keywords: cloud computing, intelligent software agent, fuzzy knowledge base, artificial neural network, cluster analysis, radical reaction, liquid phase, reactionary ability, enthalpy of reaction, bond dissociation energy

### Введение

Эффективное развитие науки и высоких технологий требует интенсивной обработки и анализа фундаментальных знаний, накопленных в различных исследовательских организациях. Это приводит к потребности в развитии информационных технологий для накопления, извлечения и анализа предметно-ориентированных профессиональных знаний на основе разрабатываемых универсальных и специализированных моделей организации и представления научных данных и знаний в электронных ресурсах для их использования профессиональным сообществом.

Таким образом, создание электронных информационных ресурсов в Интернет, предназначенных для сбора, хранения, верификации, извлечения, распространения и производства новых профессиональных знаний в различных предметных областях является актуальной научной и научно-практической задачей.

Целью настоящей работы является разработка облачной многоагентной программно-технологической архитектуры для предметно-ориентированной системы научной осведомленности по физической химии радикальных реакций [Г1] (далее система), описание назначения используемых интеллектуальных агентов в решении задач предсказания физико-химических свойств органических молекул по экспериментальным данным.

### **Краткие сведения о системе**

Подробное описание системы и принципов ее программно-технологической архитектуры дано в [5]. Система является англоязычной и включает в себя следующие информационные компоненты: базу данных по энергиям диссоциации связей органических молекул, базу данных по энтальпиям образования свободных радикалов и органических молекул, базу данных по константам скорости радикальных жидкофазных реакций.

В системе представлены активные компоненты, предназначенные для предсказания физико-химических характеристик органических молекул и их реакционной способности в радикальных реакциях (далее подсистема прогноза), а именно, экспертная система предсказания энергии диссоциации связей органических молекул по кинетическим данным, программы с элементами искусственно интеллекта для предсказания реакционной способности органических молекул в радикальных бимолекулярных реакциях и энтальпии образования свободных радикалов по кинетическим данным.

### **Методы исследования**

В процессе разработки системы на настоящем этапе были использованы следующие облачные технологии: инфраструктура как услуга (IaaS), программное обеспечение как услуга (SaaS) и рабочее место как услуга (WaaS) [2]. Кроме того, система изначально разрабатывалась в многоагентной программной архитектуре и с применением технологий хранилищ данных и знаний, интеллектуальных программных агентов, искусственных нейронных сетей, а в последнее время и технологии нечетких баз знаний.

Основными задачами исследования являются разработка и реализация программных интеллектуальных агентов для использования их в подсистеме прогноза физико-химических свойств органических молекул по экспериментальным данным для рассматриваемой системы.

### **Результаты исследования и их обсуждение**

Программно-технологическая архитектура электронного образовательного ресурса реализована на основе многоагентных технологий в частном облаке. Основные компоненты его архитектуры приведены на Рис. 1.

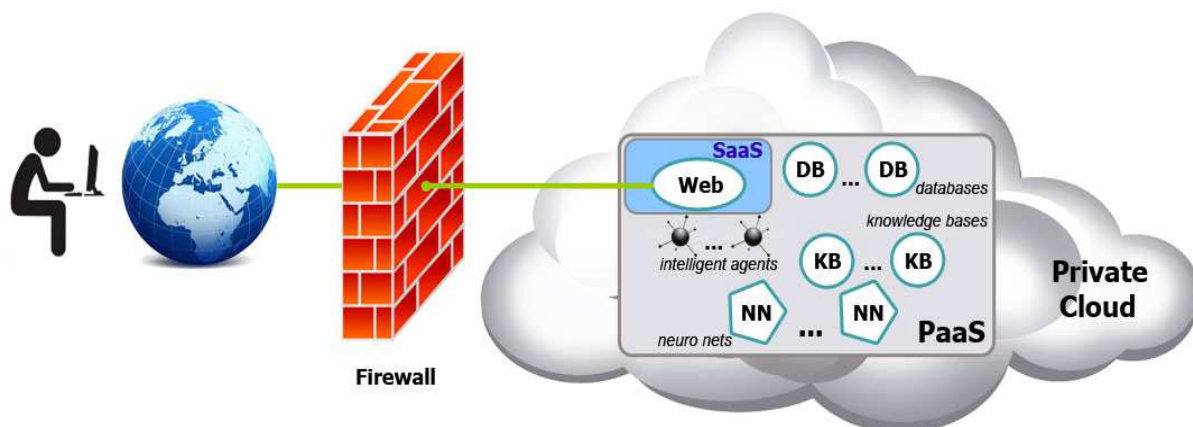


Рисунок 1. Программно-технологическая архитектура системы на настоящем этапе.

Электронные коллекции физико-химических свойств органических соединений, объединенные в системе [5], позволяют ставить и решать задачи извлечения новых знаний о свойствах этих соединений. Технология интеллектуальных программных агентов позволяет поэтапно расширять возможности разработанной системы. Анализ выборок из этих баз данных позволяет искать зависимости в данных и предсказывать новые значения физико-химических свойств органических соединений, таких как реакционная способность, прочность связи, энтальпия образования радикала.

Агентами принято называть программы, обладающие свойствами автономности, реактивности и обладающие социальным поведением в многоагентных средах [10]. Разрабатываемые нами агенты относятся к типу «Агенты по анализу и извлечению данных» [9].

Для извлечения новых знаний о физико-химических свойствах органических молекул используются три экспертные системы: для оценки энергии диссоциации связей органических молекул по кинетическим данным бимолекулярных радикальных реакций, оценке реакционной способности органических молекул в бимолекулярных радикальных реакциях по термодинамическим данным, оценке энтальпии образования свободных радикалов по кинетическим данным. Все системы используют общее хранилище знаний.

На Рис. 2. схематически показана многоагентная программная архитектура встроенных в портал экспертных систем. На приведенной схеме программные агенты функционируют в рамках простой модели «Запрос-Ответ-Соглашение». После получения входных данных производится опрос резидентных агентов. На основе полученных ответов принимается решение, какому агенту поручить выполнение предусмотренных в экспертной системе

действий. После опроса агентов формируется матрица ответов, на основе анализа которой принимается решение о том, какому агенту отдать решение задачи. При некоторых условиях решение задачи может быть отдано двум агентам.

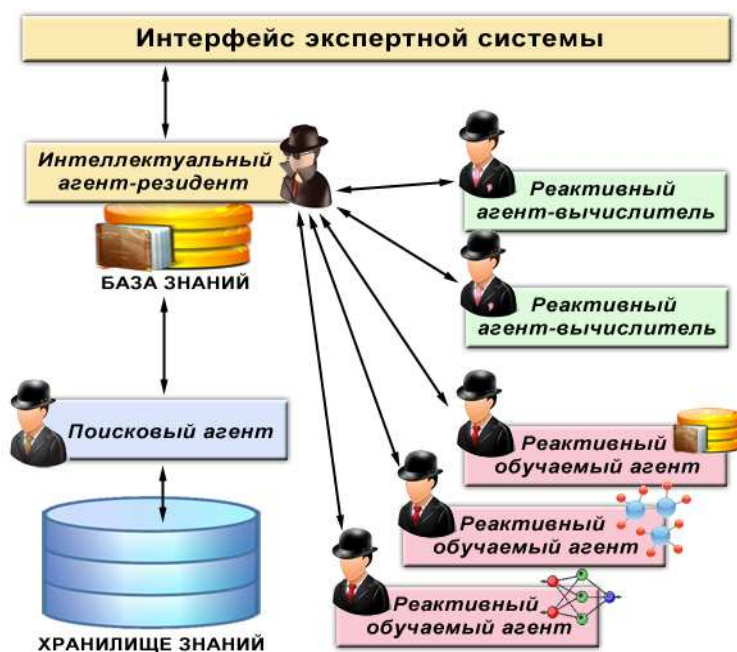


Рисунок 2. Многоагентная программная архитектура встроенной экспертной системы.

Интеллектуальные программные агенты разрабатываются либо как самостоятельные компоненты системы, либо как компоненты, которые взаимодействуют с ранее упомянутыми экспертными системами.

Программные интеллектуальные агенты, реализующие искусственные нейронные сети для предсказания реакционной способности молекул в радикальных реакциях, обученные по экспериментальным выборкам, работают как самостоятельно, так и во взаимодействии с экспертной системой оценки реакционной способности [6]. Программные интеллектуальные агенты, реализующие искусственные нейронные сети для предсказания энергии диссоциации связей органических молекул, обученные по экспериментальным выборкам, работают автономно [3]. Программные интеллектуальные агенты, реализующие алгоритмы кластерного анализа, работают автономно [7].

Программные интеллектуальные агенты, идентифицирующие зависимости в данных нечеткими базами знаний, работают автономно. Рассмотрим основы построения интеллектуального программного агента, идентифицирующего зависимости в данных. Методикам идентификации нелинейных, в том числе и эмпирических, зависимостей нечеткими базами знаний посвящено достаточно много работ, обстоятельный обзор которых можно найти в монографии [4].

Приведем пример разработки нечеткой базы знаний для идентификации эмпирической зависимости энергии активации реакций фенил радикалов ( $C_6H_5^\bullet$ , 4- $CH_3$ - $C_6H_5^\bullet$ , 4- $Br$ - $C_6H_5^\bullet$ , 4- $Cl$ - $C_6H_5^\bullet$  и т. д.) с углеводородами в жидкой фазе для предсказания последней.

Для определения пространства параметров идентификации (входных данных) энергии активации радикальной реакции  $E_{экс}$  в настоящей работе используется модель пересекающихся термов Морзе, которая определяется корреляционным соотношением [1]:

$$br_e = D_{ei}^{1/2} \ln\left(\frac{D_{ei}^{1/2}}{D_{ei}^{1/2} - E_e^{1/2}}\right) + \alpha D_{ef}^{1/2} \ln\left(\frac{D_{ef}^{1/2}}{D_{ef}^{1/2} - (E_e - \Delta H_t)^{1/2}}\right) \quad (1)$$

Параметры, входящие в формулу (1) имеют следующий смысл [Д3]:

1)  $\Delta H_e = D_i - D_f + 0,5(hLv_i - hLv_f)$  – энтальпия реакции, включающая разницу энергий нулевых колебаний рвущейся и образующейся связей. Здесь  $h$  – постоянная Планка,  $L$  – число Авогадро,  $\nu_i$  – частота колебания молекулы вдоль разрываемой связи,  $\nu_f$  – частота колебания молекулы вдоль образующейся связи,  $D_i$  – энергия диссоциации разрываемой связи,  $D_{ei} = D_i + 0,5hL\nu_i$ ,  $D_f$  – энергия диссоциации образующейся связи,  $D_{ef} = D_f + 0,5hL\nu_f$ .

2)  $E_e = E_{экс} - 0.5(hL\nu_i - RT)$  – классический потенциальный барьер активации, который включает в себя энергию нулевого колебания рвущейся связи,  $R$  – газовая постоянная,  $T$  – температура в К.

3) Параметры  $b = \pi(2\mu_i)^{1/2}\nu_i$  и  $b_f = \pi(2\mu_f)^{1/2}\nu_f$  описывают зависимость потенциальной энергии от амплитуды колебаний атомов вдоль рвущейся ( $i$ ) и образующейся ( $f$ ) валентной связи.  $2b^2$  – силовая постоянная связи,  $\mu_i$  – приведенная масса атомов для разрываемой связи,  $\mu_f$  – приведенная масса атомов для образующейся связи.

4) Параметр  $r_e$  равен суммарному растяжению рвущейся и образующейся связей в переходном состоянии.

5) Параметр  $\alpha^2$  – равен отношению силовых постоянных рвущейся и образующейся связей.

Предполагается, что классический потенциальный барьер активации задается нелинейным объектом (параметр  $\alpha$  остается постоянным для всей выборки):

$$E_e = F(D_{ef}, D_{ei}, S_1, S_2) \quad (2)$$

Для моделирование зависимости используются два качественных параметра  $S_1$  – характеристика реакционного центра углеводорода (определяемая классом соединения, рассматривается 40 классов) и  $S_2$  – характеристика растворителя (неполярный, полярный, комплексообразующий полярный, комплексообразующий неполярный). Выборка (647 реакций) получена из базы данных по константам скорости радикальных жидкофазных

реакций системы. Особенностью выборки является сочетание количественных и качественных параметров в одной модели идентификации.

При построении нечеткой базы данных входные и выходная переменные в зависимости (2) рассматриваются как лингвистические переменные, заданные на соответствующих универсальных множествах. В качестве функции принадлежности элемента к нечеткому терму  $G$  выбрана:

$$\mu^G(x) = \frac{1}{1 + \left(\frac{x-b}{c}\right)^2} \quad (3)$$

где начальные значения параметров  $b$  и  $c$  были выбраны экспертным образом. Нечеткая база знаний задается системой логических высказываний в виде:

$$\mu^{d_j}(x_1, x_2, x_3, x_4) = \bigcup_{p=1}^{k_i} \left\{ w_{jp} \left[ \bigcap_{i=1}^4 \mu^{a_i^{jp}}(x_i) \right] \right\} \quad (4)$$

Где  $x_1 \div x_4$  – лингвистические оценки входных переменных  $D_{ei}$ ,  $D_{ef}$ ,  $S_1$ ,  $S_2$ ;  $d_j$  – лингвистическая оценка выходной переменной  $E_e$ ,  $w_{jp}$  – весовая матрица правил.

Для экспериментальной выборки для реакций фенил радикалов с углеводородами нечеткая база знаний была настроена с помощью генетического алгоритма [4]. В настоящий момент она включает 634 нечетких правила.

В Табл. 1 приведено сравнение значений классического потенциального барьера активации реакции фенил радикалов с углеводородами  $E_e$ , полученными с помощью нечеткой базы знаний, с вычисленными по экспериментальным значениям энергии активации  $E_{кл\_экс}$  соответствующих реакций.

**Таблица 1.** Сравнение значений классического потенциального барьера активации реакции фенил радикалов с углеводородами.

Радикал	Углеводород	Растворитель	$E_e$ , кДж/моль	$E_{кл\_экс}$ , кДж/моль
$C_6H_5^\circ$	<i>цикло</i> -[CH=CHCH=CH(CH <sub>2</sub> ) <sub>2</sub> ]	Неполярный с комплексом	12,01	11,96
$C_6H_5^\circ$	$C_6H_5CH(CH_3)_2$	Полярный с комплексом	21,10	21,06
$C_6H_5^\circ$	$(CH_3)_2CHOH$	Полярный	21,40	21,37
$C_6H_5^\circ$	$(CH_3)_2CHOH$	Полярный	21,46	21,44
$C_6H_5^\circ$	<i>суцло</i> -[CH=CH(CH <sub>2</sub> ) <sub>4</sub> ]	Полярный	21,78	21,81
$C_6H_5^\circ$	$C_6H_5CH_2OH$	Полярный с комплексом	21,90	21,86
4-NO <sub>2</sub> - $C_6H_5^\circ$	$C_6H_5CH(CH_3)_2$	Неполярный с комплексом	24,0	23,98
$C_6H_5^\circ$	$C_6H_5CH(CH_3)_2$	Неполярный с комплексом	24,0	23,98

4-Br-C <sub>6</sub> H <sub>5</sub> <sup>o</sup>	C <sub>6</sub> H <sub>5</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	Неполярный с комплексом	24,41	24,38
C <sub>6</sub> H <sub>5</sub> <sup>o</sup>	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub> CH <sub>3</sub>	Полярный	24,43	24,42
C <sub>6</sub> H <sub>5</sub> <sup>o</sup>	CH <sub>3</sub> CH <sub>2</sub> OH	Полярный с комплексом	24,50	25,51
C <sub>6</sub> H <sub>5</sub> <sup>o</sup>	C <sub>6</sub> H <sub>5</sub> C(CH <sub>3</sub> ) <sub>2</sub> CH(O)	Неполярный с комплексом	25,58	25,55

Как видно из Табл. 1, наблюдается хорошее согласие между экспериментальными и полученными путем нечеткого логического вывода значениями классического потенциального барьера активации.

### **Выводы**

Предложена облачная программно-технологическая архитектура электронного ресурса по физической химии радикальных реакций, дано описание его функционала.

Разработанная многоагентная модель среда системы [5] позволяет выполнять анализ экспериментальных данных и извлекать новые знания и по своей архитектуре (она доступна для совместной работы пользователям интернет) представляет собой вариант системы искусственного «роевого интеллекта» (Swarm Intelligence) [8].

Впервые на примере реакций фенил радикалов с углеводородами была выполнена попытка идентифицировать зависимость классического потенциального барьера активации радикальной реакции нечеткой базой знаний, построенной на основе количественных и качественных параметров.

### **Список литературы**

1. Денисов Е.Т., Туманов В.Е. Модель переходного состояния как результат пересечения двух термов Морзе в приложении к реакциям атомарного водорода // Журнал физической химии. – 1994. – Т. 68. - № 4. – С. 719-725.
2. Клементьев И.П., Устинов В.А. Введение в облачные вычисления. – Ульяновск: УГУ, 2009. – 233 с.
3. Лазарев Д.Ю., Туманов В.Е. Экспертная система в Интернет для оценки энергий диссоциации связей органических молекул по кинетическим данным // Труды Конгресса по интеллектуальным системам и информационным технологиям «AIS-IT'09». – Том. 1. – М.: Физматлит, 2009. – С. 272-277.
4. Митюшкин Ю.М., Мокин Б.И., Ротштейн А.П. Soft Computing: идентификация закономерностей нечеткими базами знаний. – Вінниця: УНІВЕРСУМ-Вінниця, 2002. – 145 с.

5. Туманов В.Е. Система научной осведомленности по физической химии радикальных реакций / В.Е. Туманов, А.И. Прохоров, Д.Ю. Лазарев, М.Е. Соловьева // Информационные ресурсы России. – 2010. - № 5. – С. 16-21.
6. Туманов В.Е. Применение искусственных нейронных сетей для оценки реакционной способности молекул в радикальных реакциях // Информационные технологии. – 2010. - № 5. – С. 11-15.
7. Туманов В.Е. Построение групп радикальных реакций отрыва алгоритмом к-средних для предсказания физико-химических характеристик реагентов // Фундаментальные исследования. – 2013. - № 6 (часть 6). – С. 1386-1390.
8. Dzitac I., Moisiu I. Advanced AI Techniques for Web Mining. // Proceedings of the 10th WSEAS international conference on Mathematical methods, computational techniques and intelligent systems. – Corfu, Greece, 2008. – P. 343-346.
9. Haag S. Cummings M. Management Information Systems for the Information Age. — McGraw-Hill, Irwin. – 9 edition. – 2012. – 608 p.
10. Nwana H. S. Software Agents: An Overview. // Knowledge Engineering Review. – 1996. – Vol. 11. - № 3. – P. 1-40.

**Рецензенты:**

Волохов В.М., д.ф.-м.н., зав. отделом, Институт проблем химической физики РАН, г.Черноголовка.

Психа Б.Л., д.х.н., зав. лабораторией, Институт проблем химической физики РАН, г.Черноголовка.